# The sspm R package: spatial surplus production models for the management of northern shrimp fisheries

Valentin Lucet[*1] **and Eric Pedersen**[†1]

1 Concordia University, Montreal, Quebec, CA

Abstract

Productivity models such as Surplus Production Models (SPMs) models can be used to inform stock management of fisheries. However, these models often share three main flaws: (1) they are usually not spatially explicit, (2) they fail to incorporate ecosystem predictors and therefore are ill-suited to ecosystem-based management of stocks, and (3) their deployment is often limited by code availability, quality and accessibility. To fill this gap, we developed a lag-1 autoregressive spatial SPM (SSPM) based on Generalized Additive Models (GAMs), broadly applicable to spatially-structured populations, which was bundled into an R package. We applied this model to one of the most economically important invertebrate populations in Canadian waters, Northern Shrimp (Pandalus borealis) in the Newfoundland and Labrador Shelves. This stock currently lacks a population model to predict how fishing pressure and changing environmental conditions may affect future shrimp abundance in the region. Our model incorporates relevant ecosystem predictors for this stock, such as Atlantic Cod (Gadus morhua) density, alternate predator density, temperature, and stock biomass. In addition, the model is deployed through the R package sspm, a flexible framework aimed at making SSPMs easier to apply to spatially structured populations. The package allows for a repeatable and open workflow and improves the accessibility of SSPMs.

Summary

The R package sspm is designed to make spatially-explicit surplus production models (SSPM) more applicable. The package uses Generalized Additive Models (GAMs) to fit a SSPM to biomass and harvest data. The package includes a range of features to manage biomass and harvest data. Those features are organized in a stepwise workflow, whose implementation is described in more detail in Figure 2.

1. Ingestion of variables as well as spatial boundaries and discretization into patches, using the user's method of choice (random or custom sampling, voronoi tessellation or Delaunay triangulation).
2. Smoothing data using spatio-temporal GAMs smoothers.
3. Computation of productivity values taking into account harvest information.
4. Fitting of SSPMs to smoothed data with GAMs.
5. Visualization of results, including confidence and prediction intervals.
6. One step ahead prediction of biomass.

Although it was developed in a fisheries context, the package is suitable to model spatially-structured population dynamics in general.

---

[*]co-first author
[†]co-first author
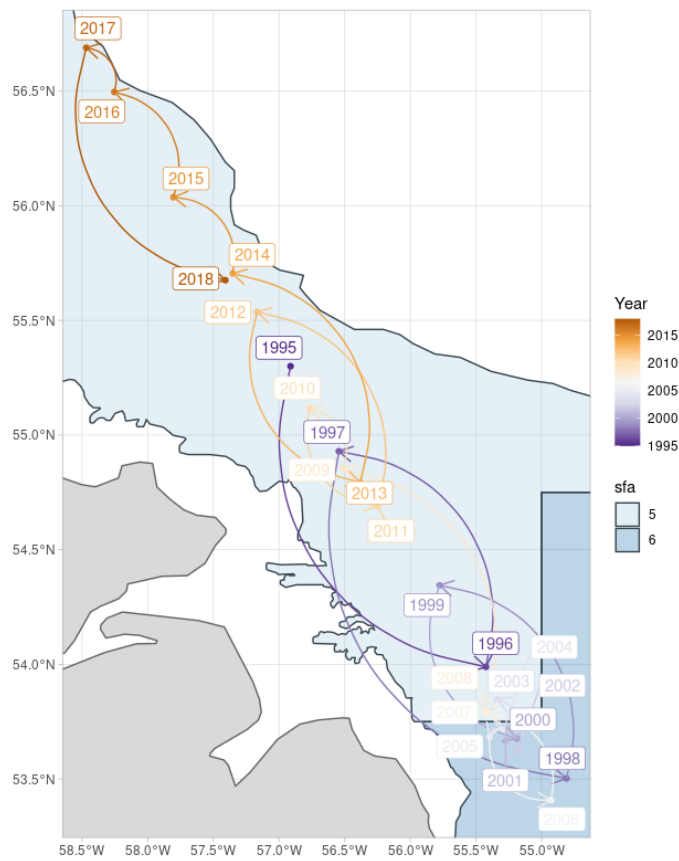
Statement of need

1. Population models, in particular fisheries productivity models, rarely integrate important spatially-structured ecosystem drivers
2. The Northern Shrimp stock in the Newfoundland and Labrador Shelves currently lacks a population model
3. Current SPM models are rarely spatially explicit and usually cannot account for relevant ecosystem drivers
4. Fisheries managers lack user-friendly, flexible tools to implement and apply SSPMs

Introduction

Population modelling is an exercise of interest within environmental sciences and adjacent fields. Early population models dating back to the blahs (80sm 90s etc, cite) tended to address simple dynamics such as exponential growth and density dependence (cite), whereas modern population models now acknowledge the non-stationary nature of wild populations. More specifically, population models applied to resource management, such as fisheries models, increasingly address how stocks vary across space and time. Resource managers are becoming increasingly interested in how ecosystem factors such as predator abundance and abiotic variables impact the spatio-temporal variability of mechanisms like productivity and density dependence. However, efforts to include spatial dynamics and ecosystem variables in fisheries models are rare. Although the non-stationarity of stocks has been established (cite), and despite the push for more ecosystem-based management methods in fisheries management (cite), applications are lacking.

One family of population models that rarely account for spatial structure is the family of Surplus production models (SPMs). SPMs are well-known tools for single-stock modelling. They model the entire biomass of a stock and are useful in data-poor contexts where the age and sex structure of the population is not accessible (cite). Basic SPMs are based on simple mechanics of logistic growth (cite), and therefore are widely viewed as a limited tool for modelling stocks. One main limitation of SPMs is that they usually assume spatially constant productivity. This assumption is a strong handicap in the context of the current global changes that are affecting global fisheries, such as climate change, which is already having an impact on the spatial structure of economically important stocks. One example is that of the Northern Shrimp (Pandalus borealis) in the Newfoundland and Labrador Shelves. This stock currently lacks a population model to predict how fishing pressure and changing environmental conditions may affect future abundance in the region. Yet the stock has experienced a consistent northward shift, as is shown in Figure 1, where the centroid of biomass trawls of the stock is plotted for each year [add estimate of speed of shift? How to calculate that?]. In this context, fisheries productivity is likely to be a moving target, and managers need better methods that account for varying productivity.

Population models like SPMs usually fall under two categories: process-based models and statistical models. Process based models often rely on differential equations and are based on replicating the underlying processes (predation, recruitment, dispersal) behind population dynamics. Statistical models, on the other hand, rely on fitting a model to data using distributional assumptions, and present the advantage of naturally measuring uncertainty around predictions. This is useful in a management context where uncertainty around decision-making is important information to have on hand. In this paper, we use a statistical approach to fitting SPMs using GAMs (generalized additive models). We apply this approach to the population of northern shrimp of the Newfoundland and Labrador Shelves, leveraging the smoothing properties of GAMs to account for varying productivity across time and space. The resulting model is a spatial SPM (SSPM), implemented via a R package sspm. We exemplify how to successfully fit a SSPM with this package and discuss the applicability of the framework to the other spatially structured populations.

Figure 1: Northward shift of weighted centroid of biomass trawled.

Model

Surplus production models (SPMs), are simply defined as such:

$$b_{t+1} = g(b_t) * e^\epsilon$$

With b the biomass, e^ϵ an error term, and g a function of the biomass and time which, in the case of fisheries, usually involves some measure of harvest through catch.

Rho => varying prod => covariance matrix => precision => basis functions and Gams

Results

The GAM biomass estimates are consistent with those of the current tool in use for the assessment of the stock, Ogmap, and provide valuable insights about the drivers of the rapid increase and decline of shrimp in the southern end of the shelf. Our approach demonstrates the model's ability to become a useful tool for modelling spatially-structured populations like fisheries stocks. The sspm package successfully modularizes each step of the modelling process and implements a range of useful features for modeling spatially-structured populations: spatial discretization, simplified GAM syntax, prediction intervals and scenario based forecasts for longer-term trends. In a fisheries context, It illustrates how our model can be easily used by managers to forecast fisheries productivity under different management regimes. The package is also a tool to think about design choices when conceiving a user interface for managers and

104 on best practices when it comes to adapting research code into management tools. Finally,
105 our approach demonstrates how open source software tools can improve the accessibility and
106 reliability of models for fisheries management.

Package design

108 The package follows an object oriented design, making use of the S4 class systems. The
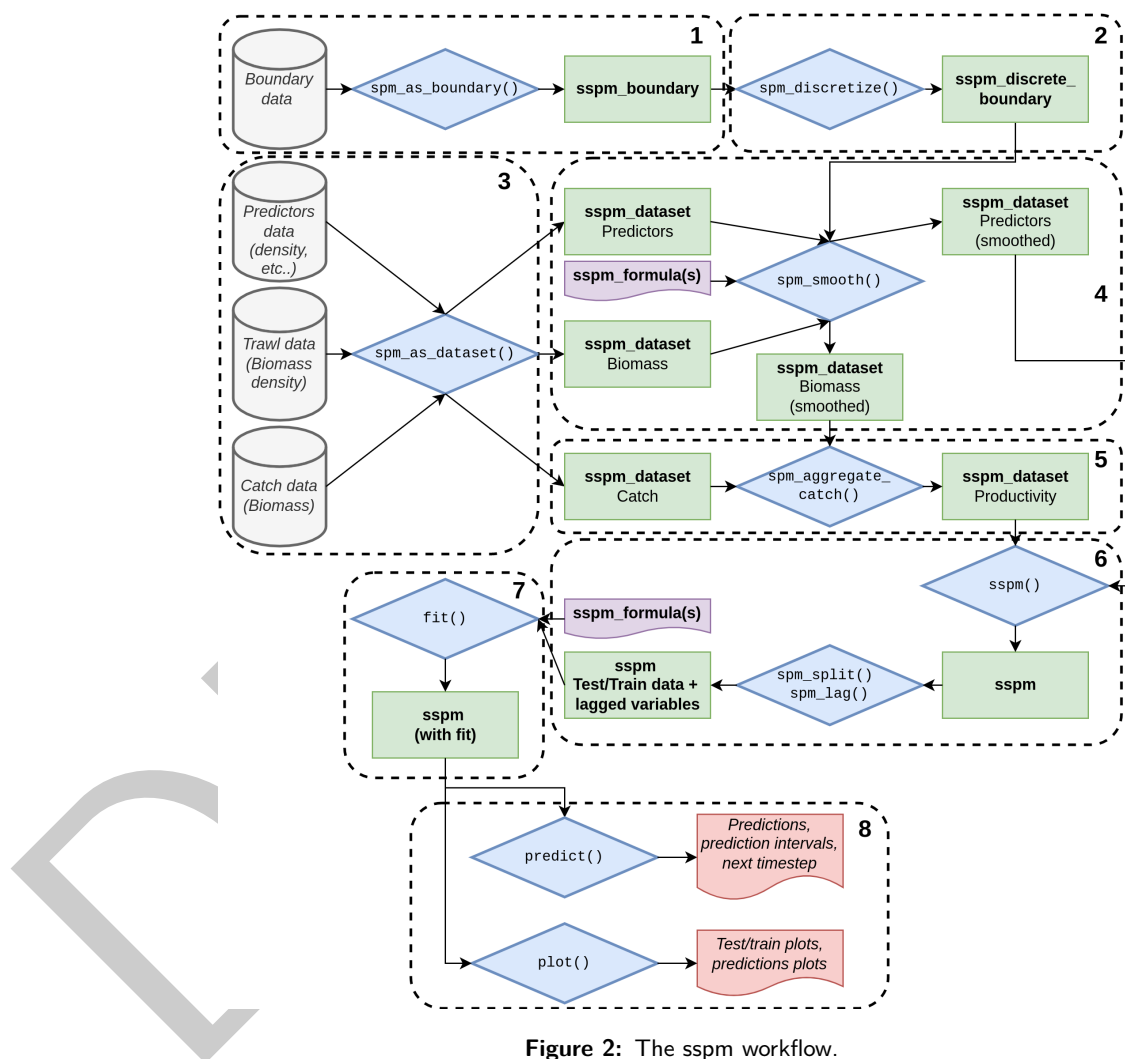109 different classes in the package work together to produce a stepwise workflow (Figure 2).

Figure 2: The sspm workflow.

110 1. The first pillar of the package's design is the concept of boundary data, the spatial
111   polygons that sets the boundary of the spatial model. The boundary data is ingested
112   into a sspm_boundary object with a call to spm_as_boundary().
113 2. The boundary data is then discretized into a sspm_discrete_boundary object with the
114   spm_discretize() function, dividing the boundary area into discrete patches.
115 3. The second pillar is the recognition of 3 types of data: **trawl**, **predictors**, and **catch**
116   (i.e. harvest). The next step in the workflow is to ingest the data into sspm_dataset
117   objects via a call to spm_as_dataset().
118 4. The first proper modelling step is to smooth the biomass and predictors data by combining
119   a sspm_dataset, and a sspm_discrete_boundary. The user specifies a gam formula with
120   custom smooth terms (see item for more details). The output is still a sspm_dataset

¹²¹ object with a `smoothed_data` slot which contains the smoothed predictions for all patches.

5. Then, catch is integrated into the biomass data by calling `spm_aggregate_catch` on the two `sspm_dataset` that contains catch and smoothed biomass. Productivity and (both log and non log) is calculated at this step.

6. The next step consists in combining all relevant datasets for the modelling of productivity (i.e. the newly created productivity dataset and the predictor(s) dataset(s)) with a call to `sspm()`. Additionally, the user may apply lags to the variables with `spm_lag()` and determine the split between testing and training data with `spm_split()`.

7. The second modelling step consists in modelling productivity per se. Once again, a gam formula with custom syntax is used (see item for more details).

8. The resulting object contains the model fit. Predictions can be obtained using the built-in `predict()` method, and plots with the `plot()` method.

[Table 1.]

Context | Syntax | Description | GAM equivalent

Context	Syntax	Description	GAM equivalent
Biomass	smooth_time()	re smoother for time steps with penalty matrix	s(bs = 're', ...)
	smooth_space()	mrf smoother for patches with penalty matrix	s(bs = 'mrf', ...)
	smooth_space_time()	ti smoother with re and mrf components with penalty matrices	ti(bs = c('re', 'mrf'), ...)
Productivity	smooth_lag()	Linear predictor smoother on a lag matrix	s(lag_matrix, ...)

Application to simulated data

We present an example using simulated biomass and harvest data. Using real trawl and fishing data provided by DFO, we generated fake data for each spatio-temporal units. …

Acknowledgements

TBD

References

TBD